CyberCog

A Synthetic Task Environment for Measuring Cyber Situation Awareness

by

Prashanth Rajivan

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Computing Studies

Approved July 2011 by the
Graduate Supervisory Committee:

John Femiani, Co-Chair
Nancy Cooke, Co-Chair
Timothy Lindquist
Kevin Gary

ARIZONA STATE UNIVERSITY

August 2011

ABSTRACT

This thesis describes a synthetic task environment, CyberCog, created for the purposes of 1) understanding and measuring individual and team situation awareness in the context of a cyber security defense task and 2) providing a context for evaluating algorithms, visualizations, and other interventions that are intended to improve cyber situation awareness. CyberCog provides an interactive environment for conducting human-in-loop experiments in which the participants of the experiment perform the tasks of a cyber security defense analyst in response to a cyber-attack scenario. CyberCog generates the necessary performance measures and interaction logs needed for measuring individual and team cyber situation awareness. Moreover, the CyberCog environment provides good experimental control for conducting effective situation awareness studies while retaining realism in the scenario and in the tasks performed.

ACKNOWLEDGMENTS

I would like to acknowledge my supervisor, Dr. Nancy J Cooke, for all of her guidance throughout my graduate studies. Her suggestions were invaluable to my thesis work. I would also like to thank the members of the Cognitive Engineering Research Institute, in particular Michael Champion and Shankaranarayan Venkatanarayanan, for their friendship, inputs, encouragement, and many interesting discussions. Finally, I would like to thank my family for their tremendous support during my university studies.

TABLE OF CONTENTS

Page

Page

Page

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

The project that is the focus of this thesis involves the development of a simulation environment and test-bed for studying cyber security situation awareness. This project is part of a larger MURI project which aims to bridge the between the gap between human and computer understanding of cyber-attack situations. This is a multi-disciplinary project involving computer science and cognitive science.

Cyber Security involves protecting the critical computing resources of an organization such as servers, software services, software applications, network connectivity and most of all, the internal information. A computer security hacker is an internal or an external agent who attempts to gain access to an organization's private network with malicious intents. The personnel trying to defend the organization from such attacks are called "cyber security defense analysts" or "computer network defense analysts" (used differently in different organizations) or simply "security analysts." The security analysts use a wide array of tools to defend the network from malicious attacks. With the growing number of cyber threats, the security analysts are overworked, cognitively overloaded and stressed, negatively impacting their work performance [1]. This makes the organization even more vulnerable to attacks. This is because the current tools and technologies intended to protect and defend our networks are still in their infancy, though the attacks have become very sophisticated and organized [2]. One of the problems with current security tools is that they do not adequately support

situation awareness (SA) (SA is defined in chapter 3). Situation awareness is necessary for operators working in complex, dynamically changing environment such as the cyber security domain. The tools used by the operators in such environments must aid the operators in achieving that necessary level of situation awareness. To build such tools, we first need to understand and measure SA in the cyber defense context. Factors found that thwart or improve SA can be identified and tools and algorithms implemented accordingly.

To study and measure situation awareness in such dynamic and complex environments, the real world tasks have to be recreated in an environment which preserves the critical complexities of the actual tasks and yet, provides the necessary experimental control. A Synthetic Task Environment [3] or an STE is one such way to achieve this.

The objective of this thesis is to collect data on cyber defense analysis and use that data to iteratively develop a synthetic task environment (CyberCog) that recreates the tasks, interaction, and team collaboration prevalent among security analysts working in a cyber security defense team. CyberCog needs to provide a rich environment for studying and measuring situation awareness in the cyber defense context. Therefore, the resulting CyberCog system needs to demonstrate that it is a suitable platform for data collection and the measurement of situation awareness.

This thesis, describes the iterative development of the CyberCog system. The first two iterations of prototype and evaluation cycles were conducted as part of a larger project. This thesis is particularly focused on refinement of the second

iteration based on information gleaned on the work of a security analyst. The objectives of this thesis are to 1) better understand the task of a cyber defense analyst, 2) incorporate this understanding in the CyberCog task at a level that non-experts will understand, and 3) embed measures of analyst performance and SA in CyberCog. The accomplishment of these objectives will be assessed by 1) demonstrating refinements in CyberCog based on insights gained from the analyst's tasks, 2) a pilot study that demonstrates that the task can be carried out by non-experts, and 3) demonstration that data are collected with potential relevance to analyst performance and SA.

The approach taken in this thesis is as follows: Subject matter experts at security analysis helped to identify the key tasks performed by cyber security defense analyst and interactions that are prevalent among the analysts within a team. Based on this information the CyberCog system was refined in its third iteration. The system includes an attack scenario and metrics to measure analyst SA and performance. A pilot study involving three participants was conducted to demonstrate that the objectives were achieved in the CyberCog system.

1.1. Problem

Cyber attacks, over the recent years, have increased exponentially in number and in sophistication. New forms of cyber threats has shifted the cyber threat landscape from simple attacks such as script kiddies to attacks for monetary purposes, to attacks for espionage and after the stuxnet [4] attack on the Iran nuclear program, the landscape has a new entry: cyber weapon. The security

analysts defending an organization from such cyber attacks are overworked and cognitively overloaded with information that is affecting their work performance and more importantly, affecting their situation awareness [11]. However, there is insufficient knowledge about cyber situation awareness to develop tools and technologies for improving situation awareness and to alleviate cognition overload in security analysts.

Governments around the globe have placed cyber security in their nation ' s top priority list [5] [6]. Private corporations perceive cyber attacks as their biggest threat. US Government networks are constantly facing a variety of cyber threats from private and state-sponsored organizations. Each day, the systems at DOD and Pentagon networks are probed and scanned hundreds and thousands of times [7]. In response, Government agencies and private corporations have been and will keep revamping their computer networks and cyber security divisions to better protect their information and infrastructure from such growing threat landscape [8]. More advanced tools, technologies and policies are being developed and are being added to the networks. More information and data are being captured and logged from the network.

Each cyber security defense analyst or simply security analysts (used interchangeably) working at cyber security divisions now monitors thousands of alerts generated from hundreds of disparate IDS (intrusion detection systems) and security sensor systems during a single work shift. Existing IDS systems are known to generate a high volume of false alerts [9] [10]. Therefore the analysts use IDS alerts as an initial source to spot any malicious activity ongoing in the

network and subsequently to initiate a cyber-attack investigation process based on the findings. Later, to prove the presence of an attack (initially identified using IDS alerts), the analysts have to peruse large amounts of raw data logs or the network packet logs. The analysts also have to review such raw data logs to identify the source of the attack, the target of the attack, vulnerabilities exploited, etc. For each alert, the analyst is given only a brief description of the attack pattern that triggered the alert and a brief description about the alert itself. Therefore cyber analysts usually rely on their experience and training to spot relevant IDS alerts corresponding to real attacks amidst a preponderance of false alarms.

The existing IDS tools and sensors use baseline network parameters to formulate an abnormal event [10] as well as known attack patterns from the attack signature database and the National Vulnerability Database to predict the current attack activity in the network and to generate an alert. With existing security tools unable to identify new threats, analysts also need to be aware of the current attack trends and the global threat level through online forums, mailing lists and intelligence reports to forecast unknown threats to their networks [9] to prepare for a future attack.

This high information and cognitive overload on security analysts greatly affects their performance and decision making abilities [1]. The present day security tools and technologies do very little to address this vital problem. Situation awareness (SA) is also challenged because there is a large amount of disparate information available to the analyst, but very little correlated

information. Also challenging is the fact that the cyber situation and attack evolve at very high speeds [11]. These unique aspects of cyber attacks such as information overload and rapid pace, along with the physical factors such as fatigue, time, pressure, illness and anxiety make the analyst's job challenging and create a critical need for security tools and visualizations that addresses the human cognitive limitations.

Cyber Security is a sociotechnical problem involving numerous individuals such as security analysts, engineers and system administrators interacting with an array of security related software tools. Therefore tools and visualizations intending to improve cyber security situation awareness in them need to be tested using a test-bed that replicates the critical complexities of the environment and facilitates human-in-the-loop experiments. Such test-beds embody ground truth (e.g., presence of an attack) and will allow developers and evaluators to get real measures of threat detection performance, thereby determining if their tools or visualizations improve situation awareness in analysts. To our knowledge, there is no such available test-bed exclusively intended for measuring situation awareness in security analyst teams.

To study situation awareness, an environment in which cyber defense experiments can be conducted with sufficient experimental control is required. However such environments are not available. Conducting SA studies in the real world cyber security defense environments or at cyber defense exercise (CDX) games are also not adequate since (1) such environments have limited ground truth availability needed to assess the performance which is an important metric in

situation awareness studies (2) such environments do not provide the experimental control needed for conducting experiments on different variables of study and across different scenarios. However these environments (actual operational environment and CDX environment) may be used to understand the work and to obtain SA requirements.

1.2. Thesis Statement

A synthetic task environment, simulating team-based cyber defense analysis work, can be developed for running empirical human-in-the-loop studies for measuring individual and team cyber security situation awareness and for testing new algorithms and tools intended for improving cyber security situation awareness.

1.3. Contributions

This research makes the following contributions:

1. Data on the work of security analyst which includes:
    a. Tasks of security analyst
    b. Interactions among security analyst in a team
2. A Synthetic Task Environment, based on an understanding of the cyber security defense context. This STE includes:
    a. A cyber-attack scenario and the corresponding datasets for the STE.
    b. Performance metrics and interaction logs for measuring situation awareness

1.4. Outline

The remainder of this document is organized as follows. Chapter 2 provides background on situation awareness, cyber security and situation awareness in the cyber security defense and analysis work. Chapter 3 presents the method adopted by CyberCog for measuring situation awareness in the cyber security defense context. Chapter 3 also presents the data obtained from running a pilot study for evaluating the system. Finally, the document concludes with a summary of this research and areas for future work.

Chapter 2

BACKGROUND

This chapter discusses background related to this research. First, a brief introduction on human factors and its importance is described. Next, definitions of situation awareness, steps to achieve situation awareness, team situation awareness and ways to measure individual and team situation awareness are described. Then, a brief description of Synthetic Task Environment is presented. Next, an overview of the cyber security domain is presented which includes the different types of cyber attacks in the threat landscape and the tools used by a security analyst. Next, the work of a security analyst is described. Finally a discussion on situation awareness in the cyber security context is presented.

2.1. Human Factors

Human Factors can be formally defined as "the study of factors and development of tools that facilitate the achievement of goals such as enhanced performance, increased safety and increased user satisfaction" [12]. It involves (1) Analyzing human interaction and tasks in a given context (2) Understanding the human capacity and limitations in performing the tasks and (3) Applying the findings towards the design and development of systems and tools

A little history of human factors which also illustrates the importance:

"After landing, pilots of Boeing B17s and B25s frequently retracted the landing gear wheels rather than the flaps. The consequence of this error was fatal. The problem was in the positioning of some key controls. The switches

controlling the flaps and the landing gear were right next to each other, and almost identical in appearance. It was not possible to move the controls further apart on aircraft that were already deployed, so Alphonse Chapanis, a pioneer in the field of industrial design, and is widely considered one of the fathers of ergonomics or human factors, modified the control by attaching a wedge shape to the control for the flaps and a small rubberized disc to the one that controlled the wheels. This eliminated so-called pilot error. This is one of the first documented incidents which gave rise to the field of human factors." [13]

2.2. Situation Awareness

Endsley defines situation awareness as "*the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future*" [14].

Situation awareness involves being aware of one's environment (i.e., being aware of all the important events and changes happening in the environment). Systems may or may not support SA and in some cases systems may actually be the cause of cognitive overload and uncertainty. Situation awareness has been studied in complex and dynamically changing sociotechnical environments such as aviation, air traffic control, military command and control, and cyber security [15]. The example that follows illustrates situation awareness in one such complex environment-driving.

Example: Bob is at a busy, unfamiliar city waiting to board a flight which departs in one hour and he also has a rental car to return. He is at the airport vicinity and is using his navigation system to find the exit for rental car returns. While looking at the navigation system he is trying to enter an exit only lane for rental car

returns. There is a vehicle in Bob's blind spot trying to enter the freeway and converging upon him. Bob is unaware of the converging vehicle and his attention is drawn to the car only when the other car honks. Due to Bob's attention towards the navigation system and due to the cognitive stress, Bob forgets to look over his shoulder while taking the exit and therefore becomes unaware of the converging vehicle. If Bob was situationally aware of all the surrounding vehicles, especially the converging vehicles on the exit only lane, he would have been more cautious while entering that lane and could have avoided the near collision.

2.2.1. Achieving Situation Awareness:

According to Endsley [14], an individual achieves situation awareness through a 3-step process as stated implicitly in the definition:

Level 1- Perception: This is the first step towards achieving SA. This involves perceiving and knowing the status of various variables and elements in an environment by monitoring the variables, cue detection and simple recognition using memory [14]. Selective attention is necessary to achieve this stage [12].

Level2 – Comprehension: The next step is to integrate all the disparate information perceived during level 1 and subsequently gain a complete understanding of the current situation. This process involves pattern recognition, evaluation and interpretation [14]. Achieving this stage involves the human memory (working and long term memory) [12].

Level3 – Projection: This is the highest level of SA achievable which involves the ability to project the future state and behavior of the elements in environment [14].

2.2.2. Measures of Situation Awareness

Situation awareness can be assessed in a variety of ways. Some of the important ways to measure SA are presented below. Techniques such as performance measures of SA, SPAM and SAGAT presented below are employed in this work.

Subjective Measures of SA:

Subjective measures of SA require individuals to rate the situation awareness they possess while performing a certain task. Participants performing real world tasks are stopped at certain intervals during the scenario and are given a questionnaire and scoring sheet to rate their current situation awareness (ex: Situation Awareness Rating Technique (SART) [16]). This approach has certain limitations in the fact that the users seldom know how much of situation awareness they possess at any point during the experiment. Halting a scenario to assess situation awareness may disrupt the subsequent task performance.

SAGAT

Situation Awareness global assessment technique or SAGAT is a technique to measure situation awareness in which the participant performing tasks in an experimental scenario is interrupted by making the interfaces they were using to go blank for a very short duration at random times. Then the

operators are asked questions related to the SA needed to perform the task. For example in a military command and control experiment the participant may be asked about the location of the nearest terrorist activity. The answers they provide are recorded and are evaluated against the ground truth to measure situation awareness [17].

Performance measures of SA:

Performance measures of SA are based on the outcome of the task. Therefore, the better the performance of the individual or a team, the better is situation awareness. Performance measures vary by experiment. Examples of performance measure are number of objectives achieved to the total number of given objectives, time taken to complete the task, the correctness of the task, etc. Performance measures are better than subjective measures because performance measures are objective and do not rely on the user to assess his or her own situation awareness. Also, by using performance measures to measure SA, the experiment need not be halted to measure or assess situation awareness. The link between situation awareness and performance is still under debate. It is argued that an individual or team may perform very well even when they have low SA [18].

SPAM

Situation present assessment method or SPAM is another technique to measure SA. It is based on the assumption that SA involves knowing where the information can be found in the environment in order to get that information and

not in remembering the information as is implied by the SAGAT blanking procedure. Opposed to SAGAT, in the SPAM method the scenario need not be halted to measure SA. Instead, the probes are present as part of the task or scenario in which task and scenario related SA questions are asked. The query response is recorded for measuring the operator's SA. For every correct query response, the time taken to respond to the query is also taken as an indicator of the operator's SA [19].

2.2.3. Team Situation Awareness:

A team is defined as a group of heterogeneous people working together towards a common goal [3]. The heterogeneity could be based on their individual skill, information they know, or their training or on the resources they have. Endsley defines team situation awareness (TSA) as "*the degree to which every team member possesses the SA required for his or her responsibilities*" [14]. The team's performance depends on the level of situation awareness in each of the team member. One member's poor SA can affect the team's performance. This view of TSA as an aggregation of individual situation awareness of members in the team can be categorized under information-processing perspective of team cognition [20]. In this view the team as a whole and individuals on the team are considered as information processors and therefore the cognition or SA measured at the individual level is aggregated to the team level [20]. This information-processing perspective however has limitations. For instance, this aggregation approach is relevant for homogeneous teams and not heterogeneous teams and

this perspective may not suffice as teams increase in size [20]. Team situation awareness is thought by some to be more than the sum of situation awareness of the individuals in the team [21]. According to this view team situation awareness emerges through the interaction of individuals in the team [21]. The team members through team interactions transform individual knowledge to collective knowledge and in the process achieve team situation awareness. Therefore, team situation awareness can be viewed as a combination of both individual's knowledge of the environment and team process behaviors such as interaction and collaboration. This view of team SA can be categorized under ecological perspective of team cognition [20]. According to this view to study team SA, a roadblock is inserted into a simulated team scenario and observations are made by monitoring the team interactions to determine (1) whether the team identifies the roadblock (2) how the team as a whole identifies the roadblock (3) what the team does to overcome the roadblock. The roadblocks are placed such that their performance will be affected if they do not identify and overcome the roadblock effectively. In this thesis team SA is examined under the ecological perspective. Therefore roadblocks need to be present in the experimental scenario and the synthetic task environment needs to elicit team interaction and collaboration.

2.3. Cyber Security

The term "cyber security" can be defined as the following:-

1. "*A set of activities and other measures, technical and non-technical, intended to protect computers, computer networks, related hardware devices and software,*

and the information they contain and communicate, including software and data, as well as other elements of cyberspace, from all threats, including threats to the national security." [22]

2. "*The degree of protection resulting from the application of these activities and measures."* [22]

3. "*The associated field of professional endeavor, including research and analysis, aimed at implementing and those activities and improving their quality."* [22]

2.3.1. Cyber Attacks

The current cyber threat landscape includes cyber attacks as simple as script kiddies to attacks such as stuxnet which is a potential cyber weapon [4]. Table 1 provides a list of all categories of cyber attacks. Most of the cyber attacks prevalent today are a combination of attack types belonging to different categories given in the Table 1. For example, an attacker could launch a buffer overflow attack on a system, exploiting an existing software vulnerability which will allow to attacker to gain administrator access or to install a rootkit. The rootkit can eventually be used to open a backdoor in the system which in turn could be used by the attacker to transfer personal information from that system to remote machines without the knowledge of the user.

Attack Type	Description
Malware and virus attack	Software designed to infiltrate a computer without the owner's informed consent intended to disrupt the user operations, for privacy exploitation, to gain system administrative access, to open backdoors, for converting the machine into a zombie computer, to retrieve system information or to perform other malicious activity on the machine.[23]
Denial of Service(DoS) and Distributed Denial of Service(DDoS) attack	Dos or DDos is an attempt to make the targeted computer resource unavailable to its intended users by maxing out the available connection threads of the resource.[23]
Botnets	It is also known as a zombie army is a group of computers operated for malicious purposes by an attacker without the owner's knowledge. These zombie computers are remotely managed by a bot master who is the attacker.[23]
Root kits	They are software that is designed to hide or obscure the fact that the system has been compromised. Root kits enable an attacker to take control of the operating system by opening a backdoor to the system. They also act to evade the operating systems security scan and antivirus software giving the user a false sense of safety.[23]
SPAM Attack	It is the flooding of internet with the same message in an attempt to force the message on the people who do not otherwise intend to receive such a mail.[23]
Software and hardware vulnerability exploitation attack	These attacks exploit the vulnerabilities of poorly programmed software which includes desktop application, web application, web services, cloud based

	services and mobile applications. Some of the most popular attacks in this category are buffer over flow attack, remote code execution, SQL injection, format string vulnerabilities, cross site scripting (XSS) and user name enumeration.[23]
Mobile and Wireless attack	They can be launched by pretending to be someone/something else such as Service Set Identifier(SSID) attacks, MAC spoofing , man in the middle attack, Wired equivalent privacy Wireless Application Protocol(WEP WAP)1.0 cracking etc. or they can result in direct denial of service attacks such as insertion attack, encryption attack and jamming.
Phishing	Phishing is an act of sending an e-mail falsely claiming to be from an established and legitimate enterprise to a user in an attempt to scam the user into surrendering private information that will be used for identity theft.
Pharming	Pharming can also be called as 'Phishing without a lure' is defined as "a scamming practice in which malicious code is installed on a personal computer or server, misdirecting users to fraudulent Web sites without their knowledge or consent"[24]

Table1: List of Cyber Attacks and description of each of the attack

2.3.2. Work of a Cyber Security Defense Analyst:

Different organizations have different job tasks and responsibilities for a security analyst. The differences are generally due to the difference in the management approach, the size of the network, the operational domain of the

organization (Ex: Microsoft or UNIX or custom environment), organizational hierarchy etc. Such differences though exist across organization, the tasks of a security analyst at a high level in any organization is the same and include the following:

- Monitoring Intrusion events,
- Collecting and filtering computer network traffic,
- Analyzing the traffic for suspicious or unexpected behavior,
- Discovering system misuse and unauthorized system access,
- Reporting to the appropriate parties
- Take actions to prevent future attacks.

Carnegie Mellon University (CMU) research classifies analyst's activities or functions into three generic groups: reactive, proactive, and security quality management [25].

- Reactive - Reactive activities are triggered by a preceding event or request such as a report of wide-spreading malicious code or an alert identified by an Intrusion Detection System (IDS) or network user complaints. Looking to the past, reactive tasks include reviewing log files, correlating alerts in search of patterns, forensic investigation following an attack and identification of an attacker who has already penetrated the network [25].
- Proactive – These activities are undertaken in anticipation of attacks or events that have not yet manifested [22]. Proactive tasks include identifying new exploits before they have been used against the defended

network, predicting future hostile actions and tuning sensors to adjust for predicted attacks [25].

- Security quality management- These activities are information technology (IT) services that support information security, but that are not directly related to a specific security event; these include security training, product evaluation, and disaster recovery planning [25].

2.3.3. Tools used by a security analyst:

The different types of tools used by a typical cyber security analyst are the following:

- Intrusion Detection Systems such as *Snort* monitors and analyzes network traffic data for any suspicious activity using pre-defined attack signatures from the signature database. The system generates alerts when network activity that matches an attack signature is detected.
- An attack signature is a unique data pattern, such as network logs, that is used by an IDS system to identify malicious network activity. Attack signature database is frequently used by an analyst to fine tune the existing attack signatures, to add new signatures or to delete a signature which may be causing too many false positives to be produced.
- Network activity and log analyzer such as *Wireshark* is used to view all the network activity in the network or view filtered activity, filtered based on IP address, time span, protocols etc. Analysts use such tools to find network activity, using filtering options, which caused the IDS system to

generate a particular alert. Analysts use such tools for further investigation after identifying interesting events using IDS tools.

- Firewall rules are modified by analysts at some companies to block or allow certain incoming network traffic.
- Patch management software is used by analysts at some companies to apply patches when the vendors make it available or when attackers exploit vulnerabilities in the unpatched systems.
- Anti-virus tools are used by the analyst to scan for virus and to quarantine the virus.
- Incident reporting tool is used by the analyst to report the presence of an attack along with findings to the forensics team for further investigation. This is usually implemented as a ticketing system wherein the analyst creates a new entry and ticket for the entry which remain open till the other responsible party closes the ticket.
- Security Management software is an end to end integrated solution of all the above listed tools usually used by analysts in large organizations.

2.4. Situation Awareness in Cyber Security Defense

Transforming data to facilitate security situation awareness, as shown in Figure 1 is one of the prime objectives of an analyst. This data transformation process begins at the raw data level. Raw network data are the most elemental data in the hierarchy as shown in Figure1. Raw data are the network traffic data at the packet level. An Intrusion Detection System (IDS) analyzes such massive

amounts of raw data using pre-defined attack signatures and when it detects network activity that matches an attack signature, it generates a Security Alert. Therefore, an IDS filters out all the normal traffic or white traffic and allows the analyst to focus only on the interesting activity happening in the network. Interesting activity is presented in the form of IDS alerts to the analyst. [25]The analyst continuously monitors IDS alerts to spot any suspicious activity which is accomplished by recognizing alerts which may pertain to an attack through their experience or by interacting with other analysts. IDS alerts are plagued with a high volume of false positives. Therefore, the analysts need to perform a triage on IDS alerts, examining the alert data and the related network activity data, to weed out the false positives. This triage process helps the analysts to narrow their focus and to conduct further investigation on a smaller set of alert data which becomes their suspicious activity set. This is the first stage of SA which involves data monitoring, attack detection and recognizing patterns that may pertain to an attack [25].

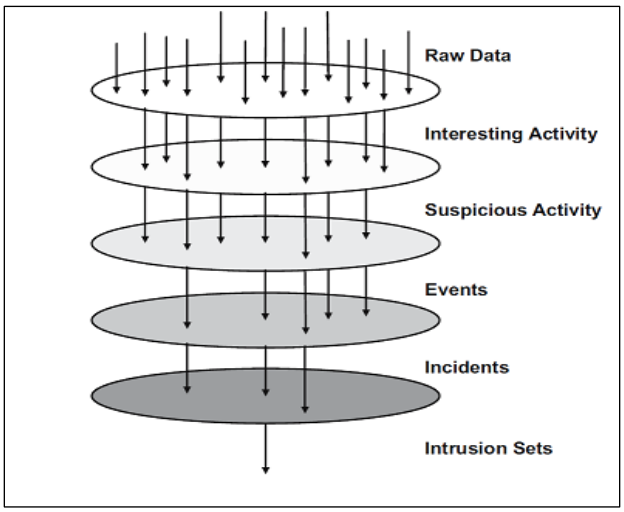

Figure 1: Raw data to situation awareness [21]

Now the analyst has to determine whether to report the suspicious activity as an event or not. Events refer to suspicious activity that an analyst has a responsibility to report, based on the organization's mission and policies and the

severity of the attack, to the concerned teams such as the forensics. For example, an organization might be obliged to report only on certain types of intrusion attempts and not on employee policy violations (e.g., using unauthorized peer-to-peer software); in this case, a policy violation would not be escalated as an event. To determine this, the analysts begin by integrating and grouping disparate suspicious activity based on certain common characteristics of the event such as the source and target IP addresses, time of the event, attack path, and attacker behavior. Along with this analysis workflow, analysts are also expanding their understanding of the suspicious activity they have by searching for and adding new facts that show the extent of the security violation including the actors, machines, and information that has been compromised. The analyst inspects the suspicious activity to gain as full an understanding as possible about that activity. If the analyst is confident that there is some attack or malicious activity taking place in the network, the analyst escalates the suspicious activity as a security event with all the findings for taking further actions. This is the second stage of SA which involves integrating data from multiple sources for gaining a complete understanding of the attack situation. [25]

The CND analyst(s), after confirming the occurrence of one or more attacks or security events, prepares a formal report describing the incident. After any required approval, the incident report is released as an official analytic product.

Intrusion sets are sets of related incidents. Intrusion sets commonly arise at the community (cyber security community) level. When a security community

suspects that separately reported incidents emanate from the same attacker source or sponsor, the community groups the incidents into an intrusion set. The community then increases attention and resources for detecting, understanding and responding to such incidents. Analysts frequently review such intrusion sets or incidents reported from different organizations. This process can include decisions about tuning data collection and IDS signatures to identify the newly found incidents in the future. This is the third and the highest level of situation awareness in cyber security defense context which involves predicting similar attacks in the future using newly created intrusion sets and attack signatures from the community. [25]

2.5. Synthetic Task Environment:

Synthetic task environments (STE) [3] are simulation environments built to recreate the real world tasks and cognitive aspects of the task in a lab environment for research studies. Synthetic task environment strives to recreate cognitive aspects of the real world task such as thoughts, distractions, analysis and cognitive overload with highest fidelity possible, giving less focus towards the appearance of the real world environment. STE differs from normal simulation environments in this aspect. Synthetic task environment provides better experimental control than uncontrolled field studies. Synthetic task environments not only provide a research environment but can also serve as a test-bed. The STE development starts by understanding and abstracting out the real world tasks and team process behaviors. This understanding is incorporated in to scenarios for

conducting experiments and studies. Then software and hardware is developed to create an interactive environment where the participants can perform the tasks within a scenario using the training provided. The software is also used to log participant actions and interactions and to collect objective and subjective measures based on the research requirements.

Chapter 3

CYBERCOG

CyberCog is a synthetic task environment that recreates the tasks, interaction and team collaboration of security analysts working on a cyber security defense team. CyberCog provides a rich environment for studying and measuring situation awareness in the cyber defense context. Individual SA and team SA are measured in CyberCog using data directly reported by CyberCog, SAGAT like knowledge elicitation methods and SPAM like probing methods are both used. Team interaction and collaboration are necessary features for measuring team situation awareness under the ecological perspective of team cognition [20] as explained in chapter 3. This chapter covers the CyberCog system including the requirements for the system, the approach taken to develop the system, a detailed description of the system and its components, and finally a discussion of the system evaluation, along with the data to support the evaluation.

3.1. High Level Requirements

The following is a list of high level requirements that should be considered while building a STE for cyber security situation awareness studies:

- Primarily, the STE (Synthetic Task Environment) should recreate the tasks performed by a Cyber Security Defense analyst.
- The STE must engage team interaction and collaboration. Observing team interactions, studies and measures on how individual SA is transformed to team SA in the cyber defense context can be made [20].

- The STE should be successful in recreating the cognitive aspects of the tasks such as information overload and analysis process found in the cyber defense context.
- Mixed Fidelity - maintaining realistic user tasks and information overload while using simplified attack scenarios and datasets. This mixed fidelity approach allows the experimenters to train participants who may not be experts in cyber security in a reasonable amount of time to perform real life tasks.
- A team by definition is heterogeneous. Each person in the team must have some distinct role or responsibilities and must have some unique tasks or information. There may be information and tasks which is common to all participants. No one person should have all the information.
- STE needs to be a distributed system enabling other MURI partners to take advantage of the system as a test-bed.

3.2. Technology

CyberCog is a distributed system built with the intention that eventually it can be used by MURI project members from other universities to test their tools and visualization. Each software interface is an ASPX page and is accessed through an Internet browser such as the Google chrome. CyberCog is developed using the ASP.NET framework and Microsoft SQL server. ASP.NET is one of the best and the widely used web application frameworks developed by Microsoft. Microsoft SQL server is a relational database from Microsoft, a preferred

database for application built using ASP.NET. Microsoft's ADO.NET entity framework is used in CyberCog which allows the application to access data tables as objects.

3.3. Approach

An iterative and incremental development process was followed for building the CyberCog system. At the end of each development phase, the system was presented to subject matter experts and real world analysts to receive feedback on the scenarios, user tasks and information presented to the participant. This feedback process helped to achieve experimental realism. In addition to expert feedback the design of CyberCog was also informed by observations of blue teams (Cyber defense team) at work during Cyber Defense exercises

We have completed 3 iterations of development and currently have a sufficient system which is ready to conduct experiments on situation awareness in cyber security analyst teams. The literature review and first two versions of CyberCog was team work while the third and the current version of CyberCog is my individual contribution. In the following section I will describe each iteration and will show how the feedback process has helped in evolving the system and how the current system is effective in measuring team and individual situation awareness in cyber security analyst teams.

3.4. First Iteration:

In the scenario developed for the first iteration of CyberCog three student participants worked together as a cyber security analyst team. Each participant

was assigned an attack based role such as Malware specialist, Denial of Service specialist and Phishing attack specialist. Each participant was trained to look for an alert pattern that is specific to his/her role. Each participant was also given a list of scenario specific cues which he/she had to use to identify the attack relevant alerts for that scenario. The cues were given to emulate the knowledge of a real world analyst. The intrusion alerts for each participant were composed of more than one type of attack. So if a participant encountered alerts which he/she was not familiar with or did not correspond to the given cues, he/she had to share it with the rest of the team to find someone who could respond to it. The team has to identify all the attack relevant alerts and submit their findings.

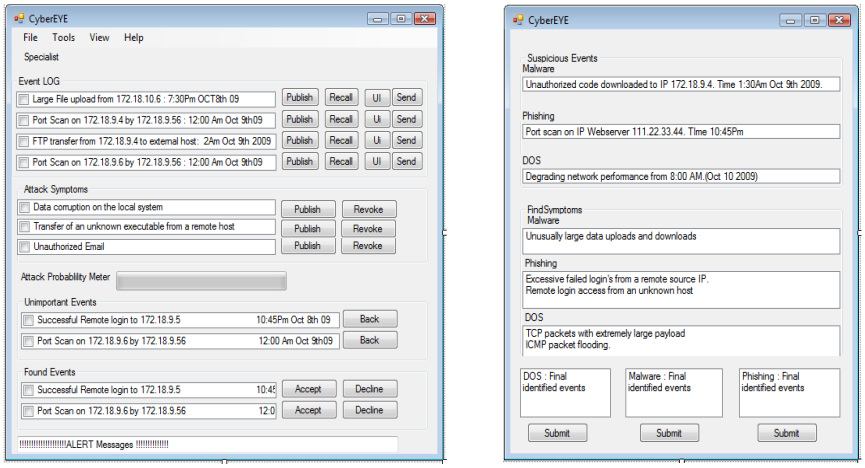

Figure 2: Snapshot of CyberCog during the first iteration

Last year we presented a working prototype of this approach, as shown in Figure 2, to subject matter experts in our MURI project team using a simple scenario and a sample dataset. We found the following:

1. There are no such specific role names in the real world cyber security analyst team.

2. The work is not divided by attack types. The division of work is *ad hoc*. The teams organize themselves and split work in an *ad hoc* manner while responding to an attack.
3. Other than intrusion alerts or events and network activity, an analyst also uses many other data sources such as system vulnerability information, systems and network information, websites and internet forums.
4. The work of the analyst team does not end at identifying attack relevant alerts but they have to identify the affected systems, vulnerabilities exploited and also have to report their findings to upper management and the forensic department who take further actions. In some cases the analysts themselves conduct attack response and mitigation duties.

3.5. Second Iteration

In the second iteration, based on the feedback obtained on the first iteration, three participants performed the roles of security analysts. Each analyst now was assigned the responsibility of a sub-network in a fictional organization. Participants received a mix of common and role-specific training. For example, all the participants were trained on the fundamentals of cyber security and on attack types, whereas each participant received some specific training on vulnerabilities in servers or system in his or her sub-network and the cues to identify if those vulnerabilities had been exploited. Thus, as in the real situation, each team member had certain specialized skills.

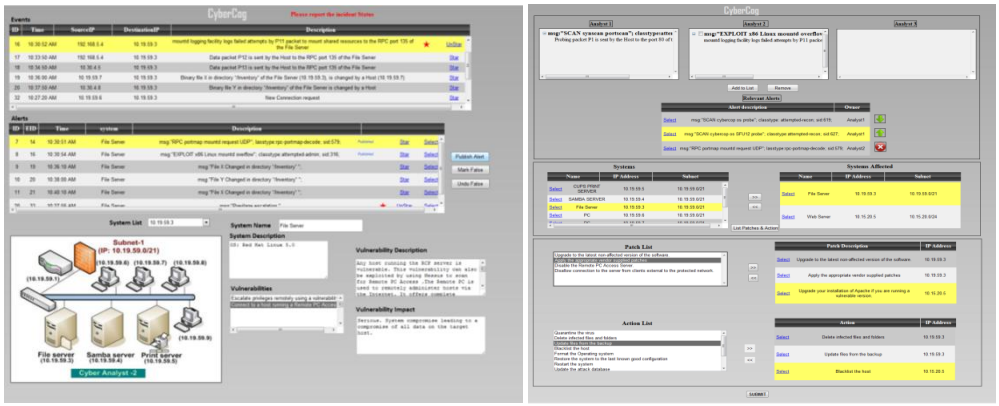

Figure 3: Snapshot of CyberCog during the second iteration

After completing the training, the team was given a scenario and a network of the organization they had to defend. Each analyst was provided with two screens of data as shown in the Figure 3. The first screen presented network activity logs and attack events from the sub- network for which the analyst was responsible. A map of the sub-network alone was provided on this screen. Each analyst, using the training provided and by interacting with the team, had to identify the relevant alerts pruning off the false ones and then sharing them with the team. The other screen helped the analyst to share his or her findings and to take response actions on the findings collaboratively. Once the team reached consensus on a certain attack event and identified the affected system, they were given the ability to select a response action which included software patches from a predefined set to mitigate the attack. The team also had to identify the attack path by sequencing the relevant alerts identified.

Each response action option had a certain cost associated with it and cost was relevant to the team performance score. The whole exercise had a time limit and the team has to defend the network within the given time. To measure situation awareness qualitatively we introduced dynamic factors into the

environment to observe how the individuals and the teams were able to identify the change and take action to mitigate the threat [26]. Factors that impacted team performance included working on false alerts, applying a patch in response to a false alert, and the time taken to apply the patch in response to an alert.

We presented this approach at the cyber situation awareness workshop we conducted in February 2011. Many real world cyber security analysts, security experts and cyber security researchers participated in this workshop in addition to the MURI members. The following are the key findings:

1. A cyber defense analyst team is usually composed of 3 to 4 security analysts.
2. Cyber security analysts work is not split by sub-networks.
3. Cyber security analysts mostly monitor the inbound and outbound traffic at the network border.
4. Network maps are seldom used.
5. Each security analyst has a different work experience, training and knowledge about the attacks and alert pattern. For example they differ by domain experience: analysts working with Microsoft based systems and analysts working with UNIX based systems. Similarly some safeguard apache based servers, whereas others safeguard IIS based servers.
6. Every day, security analysts monitor websites and online forums to keep track of new attack trends.
7. Security analysts monitor intrusion alerts from tools such as *Snort* for the most part.
8. Security analysts attend to alerts based on the priority or severity of the alert.

9. Security analysts classify incoming alerts into multiple categories such as unauthorized access, reconnaissance, false positive, denial of service etc.
10. Security analysts spot interesting or attack activity based on the alert patterns which they know pertains to an attack through experience, training, research or through patterns newly found from online forums.
11. Security analysts for the most part interact with other security analysts about the alerts they see and if they see unfamiliar patterns they tend to get help from other analysts in their team or outside their team.
12. Security analysts interact with other analyst about their findings.
13. Security analysts interact to identify new alert signatures for the novel attacks they have discovered.

3.6. Third Iteration:

The current version of the CyberCog is based on the feedback on the two previous versions as well as lessons learned from the cyber literature and cyber exercises. CyberCog is a web based application and therefore the software screens in the system are individual web pages accessed using web browsers such as Google chrome or Internet Explorer. Each participant computer is equipped with two computer monitors for running a total of six browser tab windows. Each web page corresponds to either a real world tool or a real world task. Four of the six web pages are operated using the left monitor and the remaining two web pages are operated using the right monitor. The four web pages, operated using the left monitor, are: (1) Event viewer (2) Network activity viewer (3) Classified events

viewer (4) One of these pages based on the analyst role: Network Map viewer, Software and hardware vulnerability query system, Website and Online forums. The two other web pages, operated using the right monitor, are: (1) Shared events viewer (2) Response planner. Each of the listed pages is described in detail later in this chapter.

Each participant is trained to be a cyber security analyst in this simulation environment. Participants are given a mix of common and role-specific training. Therefore each participant is equipped with some specialized knowledge distinct from other team members. The specialized knowledge is assigned to participants to emulate specialized experience that analysts acquire. Each participant receives a distinct set of events or alerts on their event viewer page. The participant has to use the training, use team's help and other information sources to investigate further on each of the given events. The investigation process begins by perusing the network activity logs to find activities which caused the alert to be raised. This is accomplished by filtering the network activity based on the event source IP and destination IP addresses and a time value. Using the network activity logs the analyst will get more information such as the source of the activity, a possible user name responsible for the activity, the reason why the alert was raised and other supporting data. The analyst may also want to use other information sources such as network map, website information, employee or user database (to know if it is an insider attack) and vulnerability information to conclude whether the event pertains to an attack or if it is a false positive. The participants have to work together as a team to find all the attack relevant alerts, to identify affected

systems, to find the attack path and to plan a response action plan to mitigate the attack.

3.6.1. Team Size:

Three participants perform the roles of security analysts. This is similar to real life security analyst teams that are typically of size three or four.

3.6.2. Training:

The participants initially undergo a training process to work as a cyber security defense analyst team using CyberCog. They are first trained on general security concepts, terminologies and technical abbreviations. The receive training on the goals of the experiment, the tasks they have to carry out during the experiment and on how to perform the tasks using the CyberCog tool. They are then trained on how to carry out a cyber-attack investigation process, a multi-step process, using CyberCog. Each participant is then trained on a distinct set of alert patterns which pertains to an attack type such as Denial of service or buffer overflow attack. They are also trained on how to investigate in the event of such an attack, what actions needs to be taken to investigate such an attack and what response actions have to be taken to mitigate such an attack. This form of training is given to emulate the background knowledge of a real world analyst and also to obtain a heterogeneous team, comparable to the heterogeneity among actual analysts.

3.6.3. Scenario

The team of analysts is given the responsibility of the network. They have to defend the network from attacks for some stipulated amount of time. During this period the team has to monitor for attacks using the intrusion alert system. They are not told if they would be attacked in the scenario. This way the participants are not actively looking for an attack from the start of the scenario and therefore a level of uncertainty is maintained. They have to classify alerts, identify if there is an ongoing attack, find affected systems and have to build an effective response action plan to mitigate the attack.

The scenario used in this version of CyberCog is a multi-step attack scenario. This scenario was constructed based on the scenario reported by Peng Liu [27]. The attacker launches a buffer overflow attack on one of the vulnerable applications running at the webserver (IP: 10.15.20.5) in the network. The map of the network used in this scenario is given in the Figure 4. Initially the attacker gets information about all running applications and services in the webserver using port scans and information retrieval queries. After launching a successful buffer overflow attack, the attacker executes a custom code in the server memory and gains administrative access on the server. On gaining administrative access on the webserver, the attacker next tries to gain access to the file server (IP: 10.15.59.3)

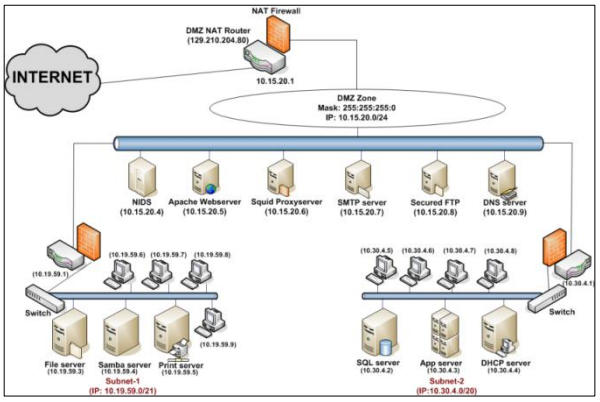

Figure 4: The network map used in the current scenario

in the same network but situated in a different sub-network. The attacker does a scan on the file server to identify all the running services and open ports in the file server. The attacker then launches a buffer overflow attack on the remote procedure call (RPC) service running in the file server exploiting an existing vulnerability. Then attacker then gains administrative access on the file server by executing a custom code on the server memory. On gaining administrative access the attacker modifies one of the files to a virus. The systems which are mounted to the file server also get affected by the virus.

3.6.4. Event viewer:

This is the primary screen which the participants would use during an experimental session. This screen simulates an Intrusion Detection System such as Snort used by analysts to monitor network intrusion alerts. The screen, as shown in the figure, displays a list of events or alerts populated in real time. Events are produced throughout the session with a particular timespan during the session being the peak time when most of the important alerts are generated. Events are unique to each analyst where the analyst 1 does not receive the same list of events as the analyst 2 and analyst 3. The events received by each analyst are of different patterns and not just the patterns on which the particular analyst received training. This ensures a certain level of uncertainty in the participant about the alert patterns. The participant may get the help of another team member regarding the unfamiliar alert patterns. For each event the user is presented with a

source IP, destination IP, time at which the event was generated and event signature which provides a one line description of the event or alert.

Key Tasks Involved:

- Alert Monitoring: The participants monitor incoming alerts or events for any anomalous activity using the training given on the alert patterns.
- Alert Starring: The participants are allowed to star events which they find interesting or which they think may pertain to an attack. This starring of events helps the participant to distinctly identify important events from the rest of the hundreds of events.

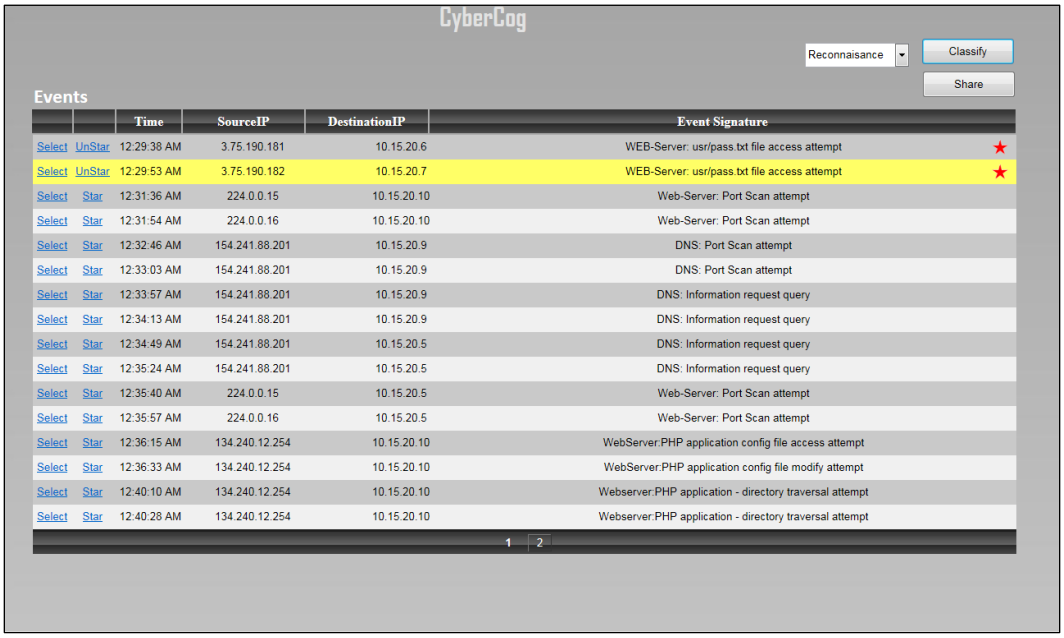

Figure 5: Snapshot of the Event viewer software screen

- Event Classification: The participants have to classify the events into one of the given categories such as false positive, denial of service,

unauthorized access etc. On classifying an event the event is removed from this list is populated in the classified events viewer page.

- Event Sharing: The participants are given the ability to share unfamiliar events with the team to get help on how to respond to such an event. This is similar to interaction among real world analysts on a team. On sharing an event, it is removed from this list and is populated in the shared events viewer page.

3.6.5. Network Data Viewer:

The network data viewer page is used by a participant to access all the network activity or to access specific activity using filtering options. The participant is presented with all of the network activity and not just the activity corresponding to the unique list of events the analyst received in the event viewer page. This is similar to the real world in which every analyst has access to all network system logs. This page simulates the real world network activity viewer or packet sniffing tool such as *Wireshark*. For each network activity the source IP address and destination IP address of the activity and a brief description about the activity (in one line) is provided. More information about the activity can be obtained by selecting the activity using the select button. The more information is displayed in the payload text box. The kind of information presented in the text box varies by the activity. This is similar to the real world tool in which the payload information of the network activity is obtained by selecting the activity from the list.

Key Task Involved:

- Data Filtering: The participants are able to filter the network activity to focus only on a small subset set of activity for which the participant is required to perform the investigation. The participants are allowed to filter by IP address and time span of the activity. The event information has IP address and time and therefore can used to narrow the search on the network activity.

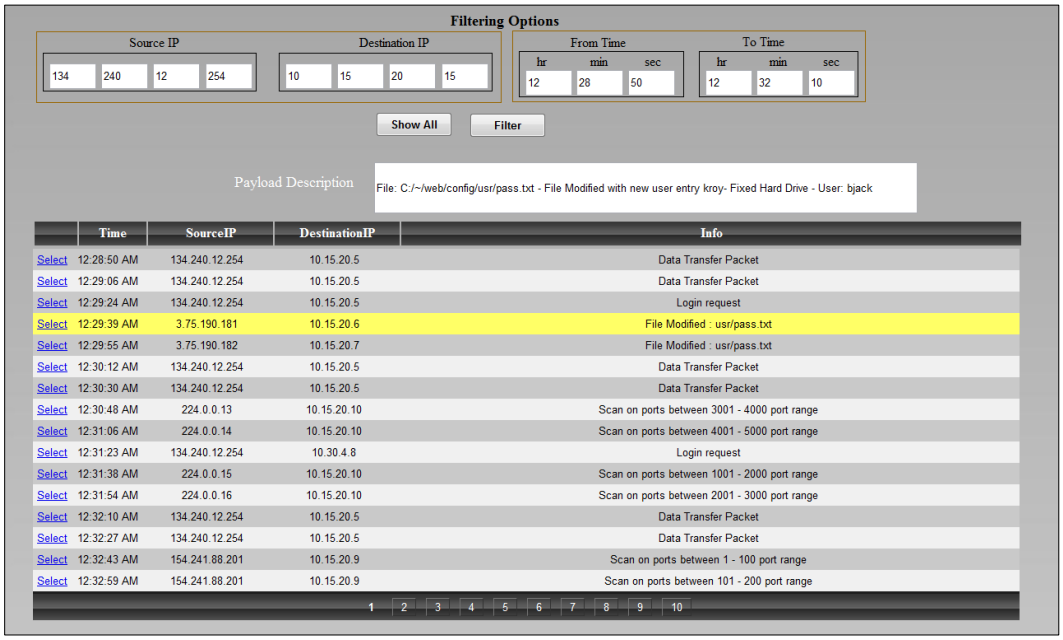

Figure 6: Snapshot of the network activity viewer

3.6.6. Classified Events Viewer:

The classified events page, as shown in figure7, is used by the participant as a placeholder to hold all the classified events. On classifying an event from the event viewer page the event is removed from the event viewer page and is populated under the corresponding category in this page. The analyst has the

ability to undo the classification and to move the event from this page to event viewer page. This helps the analyst to triage and focus on the important events. This also helps in measuring the situation awareness of the individual by assessing the accuracy of all classified events.

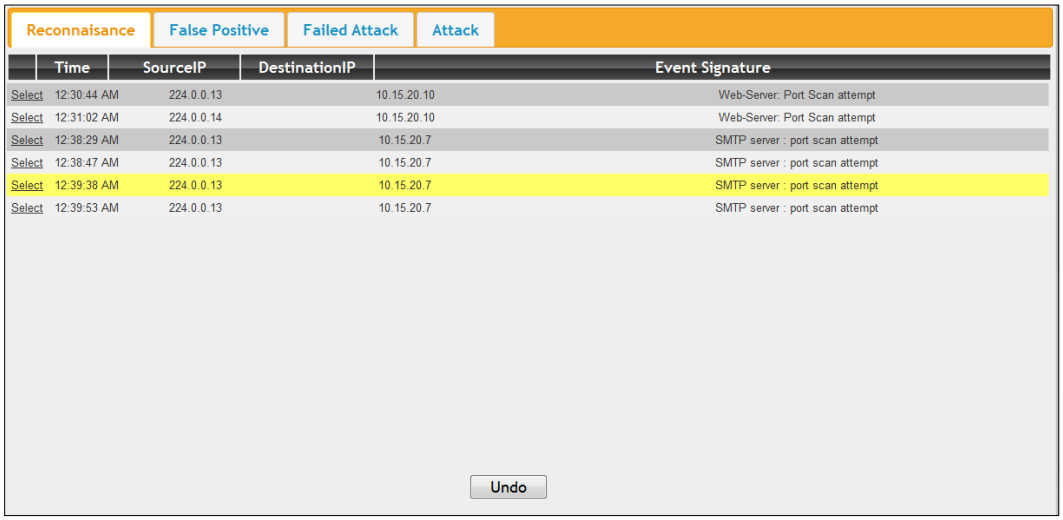

Figure 7: Snapshot of the classified events viewer

3.6.7. User Information Query system:

The participants may use this page to get information about a particular user using a username. The participant may use this to know if the user is an employee or is an unknown user, the job profile of the user, the access rights of the user, etc. This information helps the analyst to identify false users, users who gained access through malicious ways, and to spot the user during an insider attack.

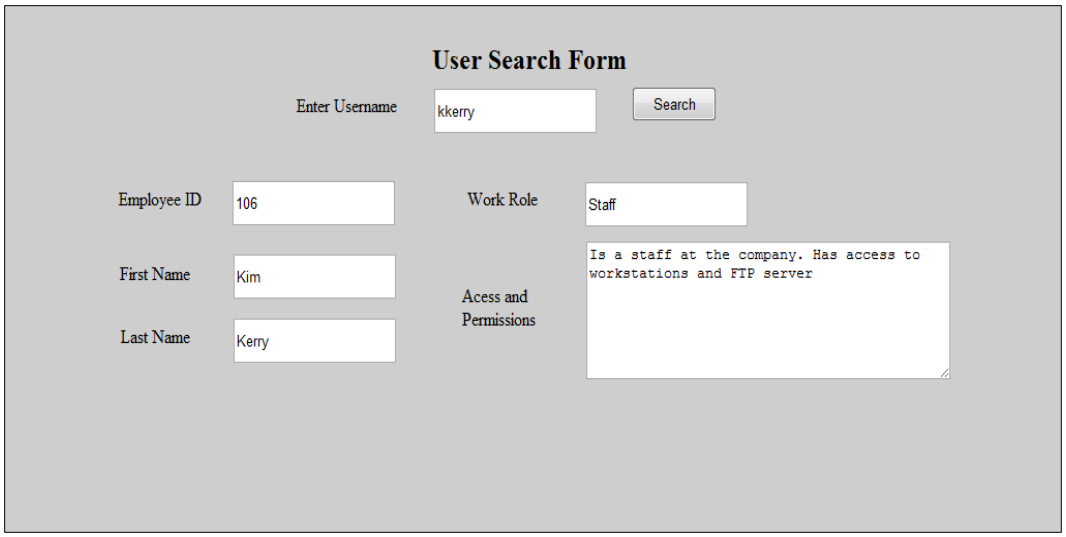

Figure 8: Snapshot of the user search form

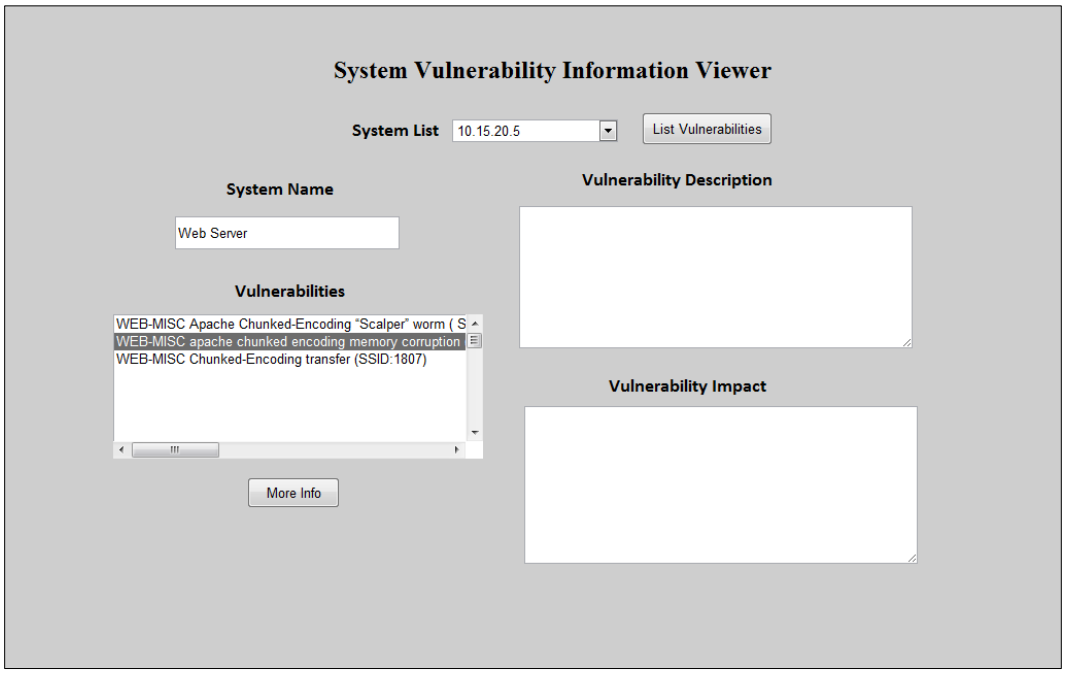

Figure 9: Snapshot of the vulnerability information viewer

3.6.8. Vulnerability Information Viewer:

The Analyst 2 participant has access to this page. The participant uses this page to find the vulnerabilities in a system and the impacts if that vulnerability gets exploited by an attacker. The participant may use this page to get more information about vulnerability in a system if the participant suspects that an attacker is trying to exploit a vulnerability in a system. Other team members have to request information Analyst 2 on vulnerability related queries. This creates one form of interaction between the participants and helps in measuring situation awareness.

3.6.9. Shared Events Viewer:

The shared events viewer page is used by the participants as a collaboration tool. On sharing an event from the event viewer page the event is removed from the event viewer page and is populated on this page. The participant has the ability to move the event from this page back to event viewer page by clicking the remove button. Only the owner of the event will have the option to remove the event. The participants share event information to get help with unfamiliar event patterns. Other team members may reply to a shared event with details and information on what needs to be done and how to carry out an investigation process for this event pattern. This interaction is very similar to interaction pattern among analysts in the real world. This interaction is therefore very crucial to measuring situation awareness. We can observe how each member

conveys his or her knowledge to other team members and how the member receiving it grasps all the information and turns it into effective actions.

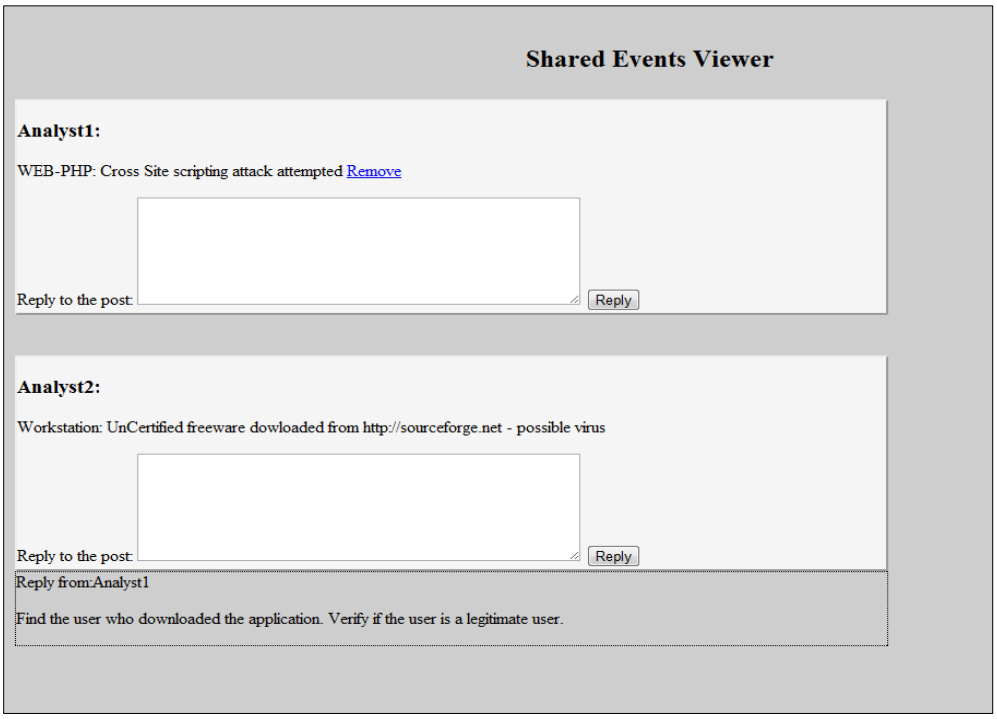

Figure 10: Snapshot of the shared events viewer

3.6.10. Response planner:

This page is used by the team as a whole to plan the response actions after identifying the attack and also to identify the attack path. The participants have to effectively collaborate to find all the affected systems and for each system they have to choose response actions to mitigate the attack on that system. Only one participant may actually use the tool at a time to plan, whereas other must collaborate with that participant to produce the plan. They may however take turns in making changes on the page. The other participants will be able to view the changes the participant is making on that page. The collaboration and

interaction taking place while doing this task is also very crucial in measuring team situation awareness. We can observe how each participant is effectively communicating his or her findings to the rest of the team and how that communication is helping the team to achieve complete situation awareness.

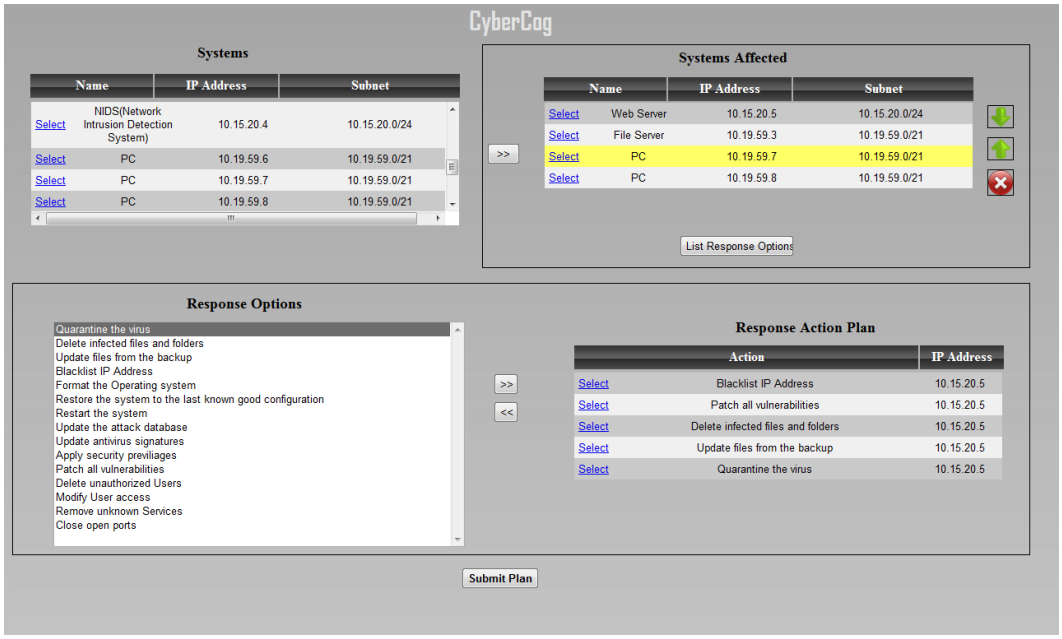

Figure 11: Snapshot of the response planner

Key Tasks Involved:

- Collaboration: Collaborate with team members to identify effected systems and to plan response actions to mitigate the attack.
- Collaboration: Collaborate with team to identify the attack path by sequencing the effected system in the order they were attacked.

3.6.11. Network Map viewer

The Analyst 1 participant has access to this page. The participant may use this page to identify the system IP address. The participant would have to use this page to understand the network connections and reachability of a system. The participant would have to use this page to identify the network path. Making it available to only team member promotes interaction among the members when they sequence the affected network systems to find the attack path and therefore helps in measuring team situation awareness.

3.6.12. Web Site Viewer

The Analyst 3 participant has access to all the websites, online forums and external intelligence. The analyst has to effectively communicate the interesting and trending attack vectors to the rest of the team cautioning the team about those attacks. This way the team becomes vigilant towards such attacks. This information will also improve their confidence in responding to those attacks if found by them.

3.6.13. Projector screen

The projector screen provides the team as a whole a dashboard of all the information such as total number of alerts, the number of classified events by category and time left to complete the scenario.

Figure 12: A snapshot of the website like form reporting the latest attack trends prevalent in the world

3.6.14. Dataset

Dataset such as network events or intrusion alerts, network activity and vulnerability information used in CyberCog are a simplified version of the real world datasets. They are simplified to plain English format removing all the technical abbreviations and number codes so that a student participant who has no experience on cyber security is able to view it, understand it and communicate the information with the rest of the team. Having a simplified dataset also helps to train the participants faster to carry out the tasks of a real world analyst. The intrusion alerts are modeled using the snort alert logs from the 2010 Westpoint CDX games [28]. For example let's consider the following real world snort alert:

[**] [1:1616:10] DNS named version attempt [**]
[Classification: Attempted Information Leak] [Priority: 2]
11/08-11:07:42.866316 10.2.195.248:50917 -> 65.190.233.37:53
UDP TTL:61 TOS:0x0 ID:52888 IpLen:20 DgmLen:58 DF
Len: 30
[Xref => http://cgi.nessus.org/plugins/dump.php3?id=10028][Xref =>
http://www.whitehats.com/info/IDS278]

The alert basically means an attacker is trying to query version and other such information of a about DNS server. So we convert this alert to this format:

DNS: Information request query
11/08-11:07:42
154.241.88.201 -> 10.15.20.5

3.7. Team Interaction

Three main types of team interactions found in actual cyber analysis and which are relevant to situation awareness studies are recreated using the CyberCog STE. The first type of interaction, as illustrated in the Figure 13, involves interaction between one analyst who is unfamiliar with certain alert patterns and another analyst on the team who recognizes the pattern and helps the analyst with details on how to investigate such a pattern. This type of interaction about attack patterns is very common in real world analyst teams. This type of interaction is realized in CyberCog by training each analyst with a different set of patterns, emulating the varied experiences of real world analysts. Analysts who are unfamiliar with a certain type of alert pattern share it with the rest of the team using the shared event viewer page and the analyst who is familiar with such a pattern responds using the same page.

Figure 13: Type 1 Interaction: Team interacts to help each other with unfamiliar event pattern

The second type of interaction, as illustrated in Figure 14, involves interaction among the analysts on the new attack patterns found at other organizations. The information is obtained by an analyst through websites or intelligence reports and is shared with the team. In this example we illustrate how an analyst's uncertainty regarding an event is resolved through such an interaction. The website-like pages provided to the analyst realizes this form of interaction.

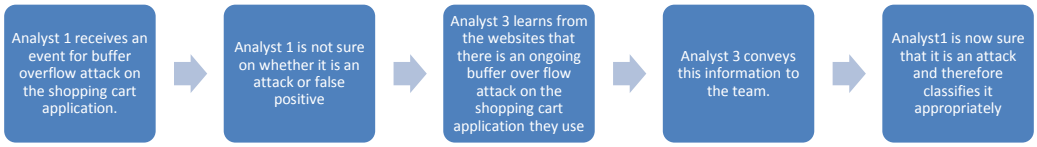

Figure 14: Type 2 Interaction: Team interacts to convey important intelligence information

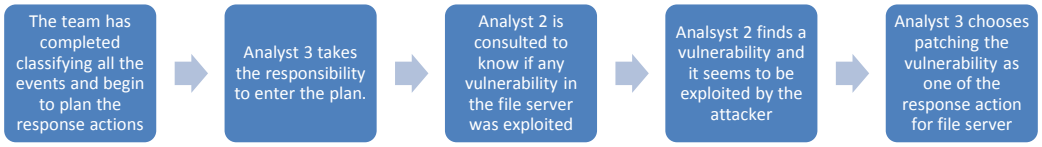

Figure 15: Type 3 Interaction: Team interacts to share the unique information they know

The third type of interaction, as shown in Figure 15, involves sharing of information or data by one analyst who has access to that resource with another analyst who is need of that information. Each analyst in CyberCog has access to a

unique source of information or resource. For instance, Analyst 2 has access to the system vulnerability information. Thus, when other analysts need to know whether a vulnerability has been exploited, they interact with analyst 2 to get that information.

3.8. SA Metrics

Situation awareness is measured in CyberCog through a combination of performance measures, team interaction analysis, knowledge elicitation method using final incident report obtained from each analyst and SPAM [19] like probing methods. Such measures helps to identify teams with high situation awareness and by looking at that team's interaction and other team process behaviors, we can identify the reasons that these teams achieve higher situation awareness than other teams. These findings will be eventually used toward the design of tools and visualization for improving situation awareness in security analyst teams.

The CyberCog STE is programmed to log all the actions performed by each analyst such as event classification, event sharing and event starring along with a timestamp to indicate when each function was performed. Each entry in the log is comma tabbed so that the output can be imported to a spreadsheet for analysis. The STE was also programmed to measure the performance of the team at the end of an experiment session.

The next section covers performance measures, user action logs and reports obtained from the pilot test and how such data can be used to ensure situation awareness in cyber security analyst teams.

Event Starred Report:

An event starred report for each analyst has a list of all the events or alerts that were marked as important by the analyst. Each participant analyst has an option to mark events which they find as suspicious activity to distinctly identify it from the rest. The report output looks like this:

```
192.121.86.47,10.15.20.10,WEB-PHP: Cross Site scripting attack attempted,6/28/2011 3:34:28 PM,6/28/2011 3:34:44 PM
```

Each entry in this report logs the event data such as the source IP, destination IP address associated with the event, the event description to identify which event was marked important. Each entry also has the time at which the event was generated and time at which the event was marked as important. The description given is in the same order as it appears in the log entry starting with source IP address. This report helps the experimenter to know if each analyst has noticed an important event and the timestamps in the entry will help in determining the time taken by the analysts to notice each event. This real time probing technique is similar to the probing technique used in the SPAM method [18].

Event Classified Report:

An event classified report for each analyst has a list of all the events or alerts that were classified by the analyst as belonging to one of the given categories. The report output looks like this:

```
224.0.0.14,10.15.20.10,Web-Server: Port Scan attempt,6/28/2011 3:20:58 PM,6/28/2011 3:48:42 PM,Recon
134.240.12.254 ,10.15.20.5,WebServer: Data received beyond the timestamp,6/28/2011 3:18:51 PM,6/28/2011 3:50:31 PM,False
```

Each log entry starts with a source IP address then destination address, event description, time the event was generated and time the event was classified and the category to which the event was classified. These logs will indicate events that were correctly identified by each analyst and the time taken by the analyst to classify it. These logs will also help identify which alerts were classified erroneously. The reason for misclassification can be learned by listening to the audio logs captured for that time period for that analyst.

Network Filter Used Report:

The network filter used report for each analyst has a list of all the filter queries used by the analyst while conducting the attack investigation. The report output looks like this:

```
135.46.574.26,10.15.20.8,6/28/2011 3:34:00 PM,6/28/2011 3:42:00 PM
10.30.4.6,132.15.623.4,6/28/2011 3:34:00 PM,6/28/2011 3:45:00 PM
```

Each log entry starts with a source IP and destination IP, address of the network activity the analyst wants within a time span given by the timestamps following the IP address in the each entry. This information highlights whether

the analyst is looking for the correct relevant network data pertaining to the attack taking place in the network which in turn is an indicator of situation awareness.

User Search Report:

A user search report for each analyst has a list of all the usernames examined by the analyst. Some network activity data provides usernames of the people initiating the traffic. The analyst examines such usernames to determine if the activity is generated by an insider or an unknown user and to know if the user performing this activity has the rights to perform the activity. The report output looks like this:

```
6/28/2011 3:44:19 PM,kkerry,Is a staff at the company. Has access to workstations and FTP server
6/28/2011 3:51:37 PM,jKing,Works on web applications. Performs changes on webserver files. Fix errors in web applications
```

Each log entry has the time at which the lookup was made, the username looked up and permissions of the user. This log indicates whether the analyst was able to differentiate activity initiated by authorized and legitimate users to the activity initiated by a malicious user.

Event Shared Report:

An event shared report for each analyst has a list of all the events or alerts that were shared by the analyst with the rest of the team for getting help in investigating unknown alert patterns. The report output looks like this:

```
162.154.26.34,10.30.4.5,Workstation: UnCertified freeware dowloaded from http://sourceforge.net - possible virus,6/28/2011 3:31:28 PM,6/28/2011 3:43:03 PM
154.241.88.201,10.15.20.5,WEB Server: Buffer Overlow attempt on the shoppingcart. Exe,6/28/2011 3:37:10 PM,6/28/2011 3:46:23 PM
```

Each entry provides the details of the event that was shared by the analyst, time at which it was generated and the time at which it was shared. The replies to the shared events are stored in the database. The database output in CSV (Comma Separate Value) format is:

1,47,Workstation: UnCertified freeware dowloaded from http://sourceforge.net - possible virus,2,3,find out who downloaded the app.. find the user name .. ifthe user is an internal user .. they u can classify it as false...and if there is no user it pertains to attack ...,2011-06-28 15:48:24.423

Each entry has some database table IDS, then the event description, then the analyst number who posted the event, then the analyst number who replied to the event and then the actual reply. The combination of these two data source captures the actual interaction (interaction type 1) between analysts on the alert patterns. For example in the given logs Analyst 2 needs help on the event "Workstation: UnCertified freeware dowloaded from http://sourceforge.net - possible virus" for which Analyst 3 responds like this "find out who downloaded the app.. find the user name .. ifthe user is an internal user .. they u can classify it as false...and if there is no user it pertains to attack …,". This is a rich source of information to analyze team interaction and for studying team situation awareness.

Final Incident Report and Confidence score:

At the end of the session after the team has submitted their findings, each analyst participant is asked to fill out an individual incident report in which they have to describe in few lines their understanding of the situation. This is a report that elicits each individual's knowledge about the situation and will help us to

determine the difference and similarity in understanding at the team level. Analyst participants are also asked to assess their confidence level regarding their work.

	Score Name	Score Description
1	Attack Events	This is the percentage of number of events correctly classified as attack events
2	False Events	This is the percentage of number of events correctly classified as false events
3	Recon Events	This is the percentage of number of events correctly classified as Reconnaissance events
4	Failed Events	This is the percentage of number of events correctly classified as failed attack events
5	Affected Systems	This is the percentage of number of systems correctly identified as affected systems
6	Response Action Plan	A correct or wrong score for the response action plan submitted.
7	Attack Path	A correct or wrong score for the attack path identified

Table 2: List of performance measures with a brief description

Team Performance Score:

At the end of each scenario a team performance score is calculated by comparing the team's findings with the correct solution. Performance measures collected from multiple teams and for multiple scenarios are good objective measures of situation awareness. The higher the performance score, the higher the situation awareness of the team. The Table 2 provides the list of all components of the performance score in CyberCog with a description of each of the score component:

Chapter 4

EVALUATION

The objectives of this thesis were to 1) better understand the task of a cyber defense analyst, 2) incorporate this understanding in the CyberCog task at a level that non-experts will understand, and 3) embed measures of analyst performance and SA in Cybercog. These objectives have been accomplished in this thesis. Supporting evidence is summarized in this section.

1) Refinements in CyberCog that are based on insights gained from the analysts' tasks have been demonstrated:

- Training each participant analyst on distinct attack patterns emulates the varying work experience and specialization of analysts in a real world team.
- Monitoring events on the Event Viewer page is very close to the way real world analyst monitor IDS alerts. The number of alerts per analyst used in CyberCog is based on the CDX games conducted by NSA in 2009 [28]
- Classifying events to different categories on the Event Viewer page is very similar to the event classification function of the analyst.
- The attack investigation process which involves the participant analyst using other information sources such as Network Data viewer, Vulnerability Information Viewer recreates the real world investigation process in a simplified manner.

- Interactions and discussions conducted by participants on the Shared Events Viewer page captures the real world interaction among analysts involving unfamiliar events and attack patterns.

2) A pilot study on a three participant team was conducted and the observations from the study demonstrate that the task can been carried out by non-experts.

The participants monitored the given events on the Event Viewer page and starred some of the events to be important. The starred event logs for each of the analysts logged such activity. The actual starred event logs from the pilot study are given below. Each log entry shows the time and date when the event was generated and time and date when the event was starred as important.

Starred event logs from the pilot study:
192.121.86.47,10.15.20.10,WEB-PHP: Cross Site scripting attack attempted,6/28/2011 3:34:28 PM,6/28/2011 3:34:44 PM
224.0.0.15,10.15.20.10,Web-Server: Port Scan attempt,6/28/2011 2:57:25 PM,6/28/2011 3:04:51 PM

Participants demonstrated that they can conduct an investigation process. Participants filtered network activity on the Network Events viewer corresponding to the events that they received. The network filter reports logged such activity. The actual logs from the pilot study are given below. Each log entry shows the IP addresses and timespan used in filtering the activity.

Network Filter used logs from the pilot study:
192.121.86.47,10.15.20.10,6/28/2011 3:34:00 PM,6/28/2011 3:38:00 PM
135.46.574.26,10.15.20.8,6/28/2011 3:34:00 PM,6/28/2011 3:42:00 PM
10.30.4.6,132.15.623.4,6/28/2011 3:34:00 PM,6/28/2011 3:45:00 PM
224.0.0.13,10.15.20.7,6/28/2011 3:25:00 PM,6/28/2011 3:45:00 PM

Participants demonstrated that they were able to classify the given events to different categories. The Event Classified reports logged the activity when each analyst classified the given event to one of the categories. The actual logs from the pilot study are given below. Each log entry shows the time and date when the event was generated and time and date when the event was classified.

Event Classified logs from the pilot study:

192.121.86.47,10.15.20.10,WEB-PHP: Cross Site scripting attack attempted,6/28/2011 3:36:37 PM,6/28/2011 3:40:30 PM,False
224.0.0.13,10.15.20.7,SMTP server : port scan attempt,6/28/2011 3:28:08 PM,6/28/2011 3:47:47 PM,Recon

Participants interacted on events that they observed. The combination of event shared reports and the replies from other analysts logged in the database demonstrates that there was interaction and that the interactions observed can be used to study team SA as per the ecological perspective of team SA [20]. The actual event shared logs and the actual replies from other analysts as observed from the pilot study are given below.

Shared Event reports from Analyst 2:

162.154.26.34,10.30.4.5,Workstation: UnCertified freeware dowloaded from http://sourceforge.net - possible virus,6/28/2011 3:31:28 PM,6/28/2011 3:43:03 PM

Reply from Analyst 3 to the above shared event:

“find out who downloaded the app.. find the user name .. ifthe user is an internal user .. they u can classify it as false...and if there is no user it pertains to attack ...,”.

3) Team performance scores collected from the pilot study demonstrate CyberCog metrics to have potential relevance to analyst performance and SA. Table 3 presents the scores collected from the pilot study. The CyberCog team performance scores include percentage scores for the events correctly categorized by the team as a whole, a percentage score for systems correctly identified as affected by the attack and a right-wrong score for the response action plan produced and attack path identified. The performance scores presented here are based on the outputs of the tasks (recreated real world analyst tasks) performed by the participants. Therefore, the scores obtained from running experiments using CyberCog can be used to measure analyst SA by using performance score approach to measuring SA. The data collected from the pilot study shows that the team on the whole was able to identify 40% of all the correct attack events, 95.8% of all the correct false events, 86.1% of all the correct reconnaissance events, 0% of all the correct failed events and 40% of all the affected systems. There was only one failed attack event in the scenario and it was wrongly classified as false positive. All of the percentage scores are calculated in the same way which is described as follows. If there are X events classified under a certain category and of the X events, Y events are the correctly identified events and if Z is the total number of correct events then the performance score is the percentage value on Y/Z. Similarly for the affected systems score, X is the number of all team identified affected systems, Y being the systems correctly identified as affected and Z being the total number of systems which are actually affected.

	Score Name	Scores from Pilot study
1	Attack Events	40%
2	False Events	95.8%
3	Recon Events	86.1%
4	Failed Events	0%
5	Affected Systems	40%
6	Response Action Plan	Wrong
7	Attack Path	Wrong

Table 3: CyberCog Performance scores obtained from the pilot study

The Response action plan and attack path are right-wrong scores obtained by evaluating the team produced plan and path against the correct plan and the actual path respectively. The teams with high SA will be able to ascertain the correct attack path.

Chapter 5

CONCLUSION

Cyber attacks growing in number and sophistication have caused cognitive overload in security analysts. The situation awareness needed for analysts to effectively conduct their tasks is impacted by this cognitive overload, as well as by ineffective tools which further overload the analysts with large amounts of disparate and false data. Security tools and technologies that aid the analyst in gaining situation awareness and tools that reduce cognition overload in analyst are needed. To build such tools that support SA, a good understanding of SA in the cyber defense context is required.

The CyberCog STE system described in this thesis is capable of providing a rich and interactive environment for studying and measuring situation awareness in the cyber defense context by recreating all of the important tasks and interaction of a real world security analyst team. Task realism was achieved by incorporating the inputs obtained through interactions with real world analysts and subject matter during the iterative development phases. Visual observations of the task obtained from observing cyber defense exercise have also helped to achieve task realism. CyberCog is also capable of recreating real world team interaction relevant for SA studies as illustrated by the three types of team interaction. The data such as the event logs, reports and performance scores obtained from the pilot test shows that the system is also capable of recording data needed to measure cyber situation awareness.

Future Work

The current CyberCog system has only one cyber-attack scenario. More scenarios along with datasets of varying difficulty levels and varying degrees of data overload should also be included in future iterations of the system to build a richer STE. Building a suite of such scenarios will allow the experimenter to conduct multiple-scenario experiments. Changing variables such as the number of attacks or number of false positives across scenarios will help in identifying the factors that affect SA in cyber analysts.

Signal Detection Theory (SDT) measures [29] such as Hit, Miss, Correct Rejection and False Alarm calculated for events identified as attack-relevant and systems identified as affected by an attack will also be included in future versions of the system. It is a "Hit" if the participant classifies an actual attack event as attack event. It is a "Miss" if the participant classifies an attack event as a false positive. It is a "Correct Rejection" if the participant classifies an actual false positive event as false positive. It is a "False Alarm" if the participant classifies a false positive event as an attack event. Similarly for the systems the team classified as affected systems, it is a "Hit" if the team classifies an actual affected system as affected. It is a "Miss" if the team does not classify an actual affected system as affected. It is a "Correct Rejection" if the team does not classify a non-affected system as affected. It is a "False Alarm" if the team classifies a non-affected system as affected.

More pilot studies will be executed to find operational gaps, to determine more cyber SA relevant measures, to improve usability of the system and also to

improve the current metrics. A complete and operational training procedure for the participants will be developed. Eventually the CyberCog STE will be used to conduct actual experiments to study and measure SA in the cyber defense context. The findings, which include factors affecting SA in security analysts, factors improving SA in security analysts and information needed to improve SA, obtained from running such experiments will be applied to designing new tools and training.

The CyberCog system contributed from this thesis work demonstrates that it satisfies the requirements of a simulation environment essential for measuring individual and team SA in the cyber defense context such as task realism, team interaction and SA relevant metrics. Therefore, CyberCog is an adequate system that can be used to conduct human-in-loop experiments to obtain cyber situation awareness measures.

REFERENCES

[1] P. Liu, "Computer-aided Human Centric Cyber Situation Awareness," 2009.

[2] J. R. Field and I. I. Operations, "CYBERSECURITY: DIVISION OF RESPONSIBILITY IN THE US GOVERNMENT," 2010.

[3] N. J. Cooke, S. M. Shope and N. Las Cruces, "Designing a synthetic task environment," *Scaled Worlds: Development, Validation, and Application,* pp. 263-278, 2004.

[4] P. K. Kerr, J. Rollins and C. A. Theohary, "The stuxnet computer worm: Harbinger of an emerging warfare capability," in 2010, .

[5] U.S. Department of Homeland Security. (2010, Open government plan. Available: http://www.archives.gov/open/open-government-plan-1.0.pdf.

[6] MINISTRY OF DEFENCE LONDON (UNITED KINGDOM), "The National Security Strategy of the United Kingdom: Security in an Interdependent World," 2008.

[7] P. W. Stewart Jim., "U.S. struggles to ward off evolving cyber threat," *Reuters,* pp. 1, 2010.

[8] G. W. Bush, *The National Strategy to Secure Cyberspace.* Morgan James Pub, 2003.

[9] A. DAmico, K. Whitley, D. Tesone, B. OBrien and E. Roth, "Achieving cyber defense situational awareness: A cognitive task analysis of information assurance analysts," in *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 2005, pp. 229-233.

[10] U. Thakar, N. Dagdee and S. Varma, "Pattern Analysis and Signature Extraction for Intrusion Attacks on Web Services," *International Journal of Network Security & its Applications (IJNSA),* pp. 190-205, 2010.

[11] K. J. Knapp, *Cyber Security and Global Information Assurance: Threat Analysis and Response Solutions.* 2009.

[12] C. D. Wickens, S. E. Gordon and Y. Liu, *An Introduction to Human Factors Engineering.* Pearson Prentice Hall Upper Saddle River, NJ, 2004.

[13] A. Chapanis, *The Chapanis Chronicles: 50 Years of Human Factors Research, Education and Design.* Aegean Pub Co, 1999.

[14] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society,* vol. 37, pp. 32-64, 1995.

[15] M. R. Endsley, "Measurement of situation awareness in dynamic systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society,* vol. 37, pp. 65-84, 1995.

[16] R. Taylor, "Situational Awareness Rating Technique(SART): The development of a tool for aircrew systems design," *AGARD, Situational Awareness in Aerospace Operations 17 p(SEE N 90-28972 23-53),* 1990.

[17] M. R. Endsley, "Situation awareness global assessment technique (SAGAT)," in *Aerospace and Electronics Conference, 1988. NAECON 1988., Proceedings of the IEEE 1988 National,* 1988, pp. 789-795 vol. 3.

[18] M. R. Endsley, "Predictive utility of an objective measure of situation awareness," in *Human Factors and Ergonomics Society Annual Meeting Proceedings,* 1990, pp. 41-45.

[19] F. T. Durso and A. R. Dattel, "SPAM: The real-time assessment of SA," *A Cognitive Approach to Situation Awareness: Theory and Application,* pp. 137-154, 2004.

[20] L. J. Rowe, N. J. Cooke and J. C. Gorman, "An Ecological Perspective on Team Cognition," 2004.

[21] N. J. Cooke, E. Salas, P. A. Kiekel and B. Bell, "Advances in measuring team cognition," *Team Cognition: Understanding the Factors that Drive Process and Performance*, pp. 83‑106, 2004.

[22] M. Dunn, "A comparative analysis of cybersecurity initiatives worldwide," in *WSIS Thematic Meeting on Cybersecurity, Geneva,* 2005, .
[23] E. Skoudis and T. Liston, "Counter Hack Reloaded: a step-by-step guide to computer attacks and effective defenses," 2005.

[24] S. Vijayalekshmi and S. A. Rabara, "Fending finanicial transaction from phishing attack," in *Trendz in Information Sciences & Computing (TISC), 2010,* pp. 171-175.

[25] A. D’ Amico and K. Whitley, "The Real Work of Computer Network Defense Analysts," *VizSEC 2007*, pp. 19-37, 2008.

[26] J. C. Gorman, N. J. Cooke and J. L. Winner, "Measuring team situation awareness in decentralized command and control environments," *Ergonomics,* vol. 49, pp. 1312-1325, 2006.
[27] P. Xie, J. H. Li, X. Ou, P. Liu and R. Levy, "Using Bayesian Networks for Cyber Security Analysis," .

[28] B. Sangster, T. O'Connor, T. Cook, R. Fanelli, E. Dean, W. J. Adams, C. Morrell and G. Conti, "Toward instrumenting network warfare competitions to generate labeled datasets," in *Proceedings of the 2nd Conference on Cyber Security Experimentation and Test*, 2009, pp. 9-9.

[29] H. Stanislaw and N. Todorov, "Calculation of signal detection theory measures," Behavior Research Methods, vol. 31, pp. 137-149, 1999.